APPENDIX 2

Propensity Score Estimation Details

Rosenbaum and Ruben (1983) defined propensity score to be the probability of treatment assignment conditional on baseline covariates ($Pr(Y_i = 1|X_i)$. The effects of measured confounding variables were removed by propensity score matching and stratification methods. Propensity score were obtained by modelling logistic regression model with statins as treatment and baseline covariates as independent variables (eq1). Sensitivity analysis was conducted by excluding intraoperative variables (eq2). Since there was missing data, multiple imputation was used to generate 10 complete datasets. The propensity score was estimated in each of these 10 datasets. The final propensity score used for matching and stratification was the mean of these 10 estimates. See the manuscript for further details on the analysis and interpretation. The following sections detail the equations used to estimate the propensity scores and standardized differences.

Eq1: Complete Covariate model

$$Pr(Statins=1|X)$$

Where X=Age, Gender, BMI, Hypertension, Diabetes Mellitus, Coronary Artery Disease, Chronic Heart Failure, Chronic Obstructive Pulmonary Heart Disease, Interstitial Lung Disease, GERD, Cirrhosis, Chemotherapy, Immunosuppression, ACE I ARB, Steroids Inhaled, Amiodarone, Antiplatelet, ASA status, Thoracic surgery, Surgical Risk Category, Anesthesia duration, Fluid balance, TV(cc/kg PBW) , Peak pressure, Smoking History, Alcohol abuse, Hemoglobin, Creatinine, and Transfusion

Table 1: Parameter Estimates from Complete Model with all Covariates

Parameter	**Estimate**	**Standard Error**	**95% Confidence Limit**
Intercept	-1.954	0.917	(-3.751, -0.156)
Age	0.015	0.006	(0.002, 0.028)
Gender	-0.125	0.074	(-0.270, 0.020)
Hypertension	-0.237	0.076	(-0.387, -0.087)
Diabetes Mellitus	-0.219	0.067	(-0.350, -0.087)
Coronary Artery Disease	-0.343	0.064	(-0.469, -0.218)
Chronic Heart Failure	0.052	0.076	(-0.097, 0.202)
Chronic Obstructive Pulmonary Disease	0.067	0.066	(-0.063, 0.197)
Interstitial Lung Disease	0.088	0.113	(-0.134, 0.311)
GERD	-0.046	0.062	(-0.168, 0.076)
Cirrhosis	0.373	0.155	(0.068, 0.677)
Chemotherapy	-0.049	0.082	(-0.209, 0.112)
Immunosuppression	-0.121	0.241	(-0.594, 0.352)
ACE I ARB	-0.171	0.062	(-0.293, -0.049)
Steroids Inhaled	-0.074	0.095	(-0.260, 0.111)
Amiodarone	-0.362	0.245	(-0.841, 0.118)
Antiplatelet	-0.589	0.060	(-0.706, -0.471)
ASA Status	-0.076	0.075	(-0.223, 0.071)
Thoracic Surgery	0.428	0.124	(0.186, 0.670)
Surgical Rick Category			
Low	0.247	0.177	(-0.101, 0.594)
Medium	0.005	0.109	(-0.209, 0.219)
Anesthesia Duration	0.001	0.001	(-0.000, 0.002)
Fluid Balance	-0.010	0.020	(-0.050, 0.029)
TV(cc/kg PBW)	-0.029	0.053	(-0.132, 0.075)
Peak Pressure	0.002	0.017	(-0.031, 0.035)
BMI	0.023	0.014	(-0.004, 0.050)
Smoking History			
Active	-0.048	0.099	(-0.242, 0.146)
Former	0.143	0.079	(-0.012, 0.297)
Alcohol Abuse	-0.051	0.096	(-0.240, 0.137)
Hemoglobin	-0.017	0.042	(-0.099, 0.065)

Parameter	Estimate	Standard Error	95% Confidence Limit
Creatinine	0.036	0.167	(-0.290, 0.363)
Transfusion	0.176	0.110	(-0.039, 0.392)

Eq2: Sensitivity Model

$$Pr(Statins=1|X^*)$$

Where X*=Age, Gender, BMI, Hypertension, Diabetes Mellitus, Coronary Artery Disease, Chronic Heart Failure, Chronic Obstructive Pulmonary Heart Disease, Interstitial Lung Disease, GERD, Cirrhosis, Chemotherapy, Immunosuppression, ACE I ARB, Steroids Inhaled, Amiodarone, Antiplatelet, ASA status, Thoracic surgery, Surgical Risk Category, Smoking History, Alcohol abuse, Hemoglobin, and Creatinine

Table 2: Parameter Estimates from Sensitivity Model Excluding Intraoperative Variables

Parameter	**Estimate**	**Standard Error**	**95% Confidence Limit**
Intercept	-1.906	0.824	(-3.522, -0.290)
Age	0.013	0.006	(0.001, 0.026)
Gender	-0.137	0.069	(-0.272, -0.002)
Hypertension	-0.234	0.076	(-0.383, -0.085)
Diabetes Mellitus	-0.220	0.067	(-0.351, -0.088)
Coronary Artery Disease	-0.338	0.064	(-0.464, -0.213)
Chronic Heart Failure	0.053	0.076	(-0.096, 0.202)
Chronic Obstructive Pulmonary Disease	0.063	0.066	(-0.066, 0.193)
Interstitial Lung Disease	0.089	0.113	(-0.132, 0.310)
GERD	-0.052	0.062	(-0.174, 0.070)
Cirrhosis	0.365	0.155	(0.062, 0.669)
Chemotherapy	-0.039	0.081	(-0.199, 0.121)
Immunosuppression	-0.130	0.242	(-0.604, 0.344)
Ace I ARB	-0.175	0.062	(-0.296, -0.054)
Steroids Inhaled	-0.075	0.094	(-0.260, 0.110)
Amiodarone	-0.355	0.243	(-0.831, 0.120)
Antiplatelet	-0.585	0.060	(-0.702, -0.467)
ASA status	-0.068	0.075	(-0.214, 0.079)
Thoracic surgery	0.290	0.096	(0.102, 0.479)
Surgical Risk Category			
Low	0.271	0.144	(-0.012, 0.555)
Medium	-0.039	0.100	(-0.236, 0.158)
BMI	0.024	0.011	(0.002, 0.046)
Smoking History			
Active	-0.041	0.098	(-0.232, 0.151)
Former	0.141	0.079	(-0.013, 0.296)
Alcohol Abuse	-0.061	0.096	(-0.249, 0.126)
Hemoglobin	-0.011	0.040	(-0.089, 0.068)
Creatinine	0.036	0.165	(-0.288, 0.359)

Propensity score stratification

The propensity scores obtained by eq1 were used to divide the subjects into 10 equal-size groups. Rosenbaum and Rubin stated that stratifying propensity scores into five equal-size groups eliminates 90% of the bias due to measured confounders when estimating a linear treatment effect. The effect of treatment on outcome within each stratum was measured by comparing the outcomes between the statin treated and untreated subjects. An overall treatment effect on outcome was measured using pooled stratum-specific estimates.

Propensity score matching

A one-to-one (greedy) matching was performed using propensity scores obtained by eq1 to select best-matched set of statin treated and untreated subjects sharing a similar value of propensity score. A caliper of 0.25 times the standard deviation of Log of propensity scores was used for one-to-one matching.

The matched set of subjects were then used to model a logistic regression model with matched pair ID as strata and Acute Lung Injury/Acute Respiratory Distress Syndrome as outcome and treatment variable Statins as independent variable to check for the association between outcome and treatment.

Standardized differences

Mean differences (eq3 and eq4) of the baseline covariates were calculated for the all statin treated and untreated subjects. Mean differences of baseline covariates were once again calculated for the one-to-one matched set of subjects. The before and after absolute mean differences were plotted to visualize the differences in the means after adjusting for treatment effect. Standardized differences less than 10% were considered negligible.

Eq3: for Continuous covariates

$$\text{Mean Difference} = \frac{\text{Mean of Statin treated subjects} - \text{Mean of untreated subjects}}{\sqrt{\frac{(\text{STD of Statin treated subjects}^2 + \text{STD of untreated subjects}^2)}{2}}}$$

Eq4: for Binary covariates

$$\text{Mean Difference} = \frac{\text{Mean of Statin treated subjects} - \text{Mean of untreated subjects}}{\sqrt{\frac{(\text{Mean}_{\text{st}} + \text{Mean}_{\text{nst}})}{2}}}$$

Where

$$Mean_{st} = [\text{Mean of Statin treated subjects} * (1 - \text{Mean of Statin treated subjects})]$$

$$Mean_{nst} = [\text{Mean of untreated subjects} * (1 - \text{Mean of untreated subjects})]$$